

## Claims

What is claimed is:

1. A method of providing a checkpoint/restart facility across a plurality of plurality of computer systems, wherein:  
the plurality of computer systems comprises:  
a first computer system executing a first program, and  
a second computer system containing a disk system and  
executing a second program;  
the first computer system and the second computer system are  
heterogeneous computer systems;  
said method comprising:  
A) checkpointing a current status of the first program resulting in a  
first set of checkpoint status information;  
B) transmitting a first checkpoint request that includes the first set of  
checkpoint status information from the first program over a first  
session to the second program;  
C) checkpointing the second program resulting in a second set of  
checkpoint status information in response to receiving the first  
checkpoint request;  
D) writing the first set of checkpoint status information and the second  
set of checkpoint status information to a first checkpoint file on  
the disk system; and  
E) transmitting a first checkpoint response from the second program  
over the first session to the first program after the writing in  
step (D) is complete.

1 2. The method in claim 1 wherein:  
2 the method further comprises:  
3 F) checkpointing the first program resulting in a third set of  
4 checkpoint status information;  
5 G) transmitting a second checkpoint request that includes the third set  
6 of checkpoint status information from the first program over the  
7 first session to the second program;  
8 H) checkpointing the second program resulting in a fourth set of  
9 checkpoint status information in response to receiving the first  
10 checkpoint request transmitted in step (G);  
11 I) writing the third set of checkpoint status information and the fourth  
12 set of checkpoint status information to a second checkpoint file  
13 on the disk system; and  
14 J) transmitting a second checkpoint response from the second  
15 program over the first session to the first program after the  
16 writing in step (I) is complete.

1 3. The method in claim 2 which further comprises:  
2 J) transmitting a first rollback request from the first program over the  
3 first session to the second program;  
4 K) reading the third set of checkpoint status information and the  
5 fourth set of checkpoint status information from the second  
6 checkpoint file in response to receiving the first rollback  
7 request transmitted in step (J);  
8 L) rolling back the second program utilizing the fourth set of  
9 checkpoint status information read in step (K);  
10 M) transmitting a first rollback response from the second program  
11 over the first session to the first program that includes the third  
12 set of checkpoint status information read in step (K); and  
13 N) rolling back the first program utilizing the third set of checkpoint  
14 status information in response to receiving the first rollback  
15 response in step (M).

1 4. The method in claim 2 wherein:  
2 the first checkpoint file and the second checkpoint file are a same file.

5. The method in claim 1 which further comprises:
- F) transmitting a first rollback request from the first program over the first session to the second program;
  - G) reading the first set of checkpoint status information and the second set of checkpoint status information from the first checkpoint file in response to receiving the first rollback request transmitted in step (F);
  - H) rolling back the second program utilizing the second set of checkpoint status information read in step (G);
  - I) transmitting a first rollback response from the second program over the first session to the first program that includes the first set of checkpoint status information read in step (G);
  - J) rolling back the first program utilizing the first set of checkpoint status information in response to receiving the first rollback response in step (I).
6. The method in claim 1 which further comprises:
- F) transmitting a second checkpoint request that includes the first set of checkpoint status information from the first program over a second session to a third program executing in a third computer system;
  - G) checkpointing the third program resulting in a fourth set of checkpoint status information in response to receiving the second checkpoint request;
  - H) writing the first set of checkpoint status information and the fourth set of checkpoint status information to a second checkpoint file; and
  - I) transmitting a second checkpoint response from the third program over the second session to the first program after the writing in step (H) is complete.

- 1    7.    The method in claim 6 which further comprises:  
2        J) transmitting a first rollback request from the program over the first  
3           session to the second program;  
4        K) reading the first set of checkpoint status information and the  
5           second set of checkpoint status information from the first  
6           checkpoint file in response to receiving the first rollback  
7           request transmitted in step (J);  
8        L) rolling back the second program utilizing the second set of  
9           checkpoint status information read in step (K);  
10       M) transmitting a first rollback response from the second program  
11           over the first session to the first program that includes the first  
12           set of checkpoint status information read in step (K); and  
13       N) rolling back the first program utilizing the first set of checkpoint  
14           status information in response to receiving the first rollback  
15           response transmitted in step (M).

9. The method in claim 1 wherein:  
there are plurality of sessions open between the first program and the  
second program for accessing a corresponding plurality of files  
by the second program; and  
the checkpointing in step (C) flushes all of the plurality of files and  
includes checkpoint information for all of the plurality of files  
in the second set of checkpoint information.

- 1 10. A computer readable Non-Volatile Storage Medium encoded with  
2 software for providing a checkpoint/restart facility across a plurality  
3 of plurality of computer systems, wherein:  
4 the plurality of computer systems comprises:  
5 a first computer system executing a first program, and  
6 a second computer system containing a disk system and  
7 executing a second program;  
8 the first computer system and the second computer system are  
9 heterogeneous computer systems;  
10 said software comprising:  
11 A) a set of computer instructions for checkpointing a current status of  
12 the first program resulting in a first set of checkpoint status  
13 information;  
14 B) a set of computer instructions for transmitting a first checkpoint  
15 request that includes the first set of checkpoint status  
16 information from the first program over a first session to the  
17 second program;  
18 C) a set of computer instructions for checkpointing the second  
19 program resulting in a second set of checkpoint status  
20 information in response to receiving the first checkpoint  
21 request;  
22 D) a set of computer instructions for writing the first set of checkpoint  
23 status information and the second set of checkpoint status  
24 information to a first checkpoint file on the disk system; and  
25 E) a set of computer instructions for transmitting a first checkpoint  
26 response from the second program over the first session to the  
27 first program after the writing in set (D) is complete.

- 1 11. A data processing system having software stored in a set of Computer  
2 Software Storage Media for providing a checkpoint/restart facility  
3 across a plurality of plurality of computer systems, wherein:  
4 the data processing system comprises the plurality of computer  
5 systems;  
6 the plurality of computer systems comprises:  
7 a first computer system executing a first program, and  
8 a second computer system containing a disk system and  
9 executing a second program;  
10 the first computer system and the second computer system are  
11 heterogeneous computer systems;  
12 said software comprising:  
13 A) a set of computer instructions for checkpointing a current status of  
14 the first program resulting in a first set of checkpoint status  
15 information;  
16 B) a set of computer instructions for transmitting a first checkpoint  
17 request that includes the first set of checkpoint status  
18 information from the first program over a first session to the  
19 second program;  
20 C) a set of computer instructions for checkpointing the second  
21 program resulting in a second set of checkpoint status  
22 information in response to receiving the first checkpoint  
23 request;  
24 D) a set of computer instructions for writing the first set of checkpoint  
25 status information and the second set of checkpoint status  
26 information to a first checkpoint file on the disk system; and  
27 E) a set of computer instructions for transmitting a first checkpoint  
28 response from the second program over the first session to the  
29 first program after the writing in set (D) is complete.

- 1 12. The software in claim 11 wherein:  
2 the software further comprises:  
3 F) a set of computer instructions for checkpointing the first program  
4 resulting in a third set of checkpoint status information;  
5 G) a set of computer instructions for transmitting a second checkpoint  
6 request that includes the third set of checkpoint status  
7 information from the first program over the first session to the  
8 second program;  
9 H) a set of computer instructions for checkpointing the second  
10 program resulting in a fourth set of checkpoint status  
11 information in response to receiving the first checkpoint request  
12 transmitted in set (G);  
13 I) a set of computer instructions for writing the third set of  
14 checkpoint status information and the fourth set of checkpoint  
15 status information to a second checkpoint file on the disk  
16 system; and  
17 J) a set of computer instructions for transmitting a second checkpoint  
18 response from the second program over the first session to the  
19 first program after the writing in set (I) is complete.



- 1 13. The software in claim 12 which further comprises:  
2 J) a set of computer instructions for transmitting a first rollback  
3 request from the first program over the first session to the  
4 second program;  
5 K) a set of computer instructions for reading the third set of  
6 checkpoint status information and the fourth set of checkpoint  
7 status information from the second checkpoint file in response  
8 to receiving the first rollback request transmitted in set (J);  
9 L) a set of computer instructions for rolling back the second program  
10 utilizing the fourth set of checkpoint status information read in  
11 set (K);  
12 M) a set of computer instructions for transmitting a first rollback  
13 response from the second program over the first session to the  
14 first program that includes the third set of checkpoint status  
15 information read in set (K); and  
16 N) a set of computer instructions for rolling back the first program  
17 utilizing the third set of checkpoint status information in  
18 response to receiving the first rollback response in set (M).
- 1 14. The software in claim 12 wherein:  
2 the first checkpoint file and the second checkpoint file are a same file.

15. The software in claim 11 which further comprises:
- F) a set of computer instructions for transmitting a first rollback request from the first program over the first session to the second program;
  - G) a set of computer instructions for reading the first set of checkpoint status information and the second set of checkpoint status information from the first checkpoint file in response to receiving the first rollback request transmitted in set (F);
  - H) a set of computer instructions for rolling back the second program utilizing the second set of checkpoint status information read in set (G);
  - I) a set of computer instructions for transmitting a first rollback response from the second program over the first session to the first program that includes the first set of checkpoint status information read in set (G);
  - J) a set of computer instructions for rolling back the first program utilizing the first set of checkpoint status information in response to receiving the first rollback response in set (I).

16. The software in claim 11 which further comprises:
- F) a set of computer instructions for transmitting a second checkpoint request that includes the first set of checkpoint status information from the first program over a second session to a third program executing in a third computer system;
  - G) a set of computer instructions for checkpointing the third program resulting in a fourth set of checkpoint status information in response to receiving the second checkpoint request;
  - H) a set of computer instructions for writing the first set of checkpoint status information and the fourth set of checkpoint status information to a second checkpoint file; and
  - I) a set of computer instructions for transmitting a second checkpoint response from the third program over the second session to the first program after the writing in set (H) is complete.



- 1 18. The software in claim 16 which further comprises:
  - 2 J) a set of computer instructions for transmitting a first rollback
  - 3 request from the program over the first session to the second
  - 4 program;
  - 5 K) a set of computer instructions for reading the first set of checkpoint
  - 6 status information and the second set of checkpoint status
  - 7 information from the first checkpoint file in response to
  - 8 receiving the first rollback request transmitted in set (J);
  - 9 L) a set of computer instructions for rolling back the second program
  - 10 utilizing the second set of checkpoint status information read in
  - 11 set (K);
  - 12 M) a set of computer instructions for transmitting a first rollback
  - 13 response from the second program over the first session to the
  - 14 first program that includes the first set of checkpoint status
  - 15 information read in set (K);
  - 16 O) a set of computer instructions for transmitting a second rollback
  - 17 request from the first program over the second session to the
  - 18 third program;
  - 19 P) a set of computer instructions for reading the first set of checkpoint
  - 20 status information and the fourth set of checkpoint status
  - 21 information from the second checkpoint file in response to
  - 22 receiving the second rollback request transmitted in set (O);
  - 23 Q) a set of computer instructions for rolling back the third program
  - 24 utilizing the fourth set of checkpoint status information read in
  - 25 set (P);
  - 26 R) a set of computer instructions for transmitting a second rollback
  - 27 response from the third program over the second session to the
  - 28 first program that includes the first set of checkpoint status
  - 29 information read in set (P); and
  - 30 S) a set of computer instructions for rolling back the first program
  - 31 utilizing the first set of checkpoint status information in
  - 32 response to receiving the first rollback response transmitted in
  - 33 set (M) and the second rollback response transmitted in set (R).

1 19. The software in claim 11 wherein:  
2 there are plurality of sessions open between the first program and the  
3 second program for accessing a corresponding plurality of files  
4 by the second program; and  
5 the checkpointing in set (C) flushes all of the plurality of files and  
6 includes checkpoint information for all of the plurality of files  
7 in the second set of checkpoint information.

1 20. A data processing system having software stored in a set of Computer  
2 Software Storage Media for providing a checkpoint/restart facility  
3 across a plurality of plurality of computer systems, wherein:  
4 the data processing system comprises the plurality of computer  
5 systems;  
6 the plurality of computer systems comprises:  
7 a first computer system executing a first program, and  
8 a second computer system containing a disk system and  
9 executing a second program;  
10 the first computer system and the second computer system are  
11 heterogeneous computer systems;  
12 said software comprising:  
13 A) means for checkpointing a current status of the first program  
14 resulting in a first set of checkpoint status information;  
15 B) means for transmitting a first checkpoint request that includes the  
16 first set of checkpoint status information from the first program  
17 over a first session to the second program;  
18 C) means for checkpointing the second program resulting in a second  
19 set of checkpoint status information in response to receiving the  
20 first checkpoint request;  
21 D) means for writing the first set of checkpoint status information and  
22 the second set of checkpoint status information to a first  
23 checkpoint file on the disk system; and  
24 E) means for transmitting a first checkpoint response from the second  
25 program over the first session to the first program after the  
26 writing in set (D) is complete.